

Appl. No.: 09/824,844
Exhibit D (Page1 of 6)

Catalyst Layer2 Trace
Functional And Design Specification

Appl. No.: 09/824,844 Exhibit D (Page 2 of 6)
--

1.0 Overview

Traceroute is a widely available utility on most operating systems and also on the Catalyst switch. It lists all the intermediate routing devices between two systems. This provides the user with the logical path of all the IP devices that a packet will traverse when going from the source to the destination. Traceroute operates by sending out a series of IP packets with successively increasing Time-To-Live counter values. When the TTL value is decremented and expires, the node at which the expiration occurs sends an ICMP message back to the source reporting the node's IP address. Thus by sending out a series of probe IP packets, the source can map the path to the destination, hop by hop. This works at the Layer 3 level - listing all the Layer 3 devices between the source and destination. In actuality, the path between any two Layer 3 entities may traverse entire networks of devices operating at Layer 2. Traceroute utility does not provide any information about the Layer 2 level devices that the packet traverses. In networks where the switched traffic can flow from the source to the destination without traversing a Layer 3 device in between, traceroute does not provide any information about the intermediate devices. In effect, traceroute is not effective in networks consisting of only Layer 2 switches/devices.

2.0 Problem Definition

The goal of this "l2trace" feature (Layer 2 trace feature) is to provide the network operator with the Level 2, or the physical path of unicast traffic, i.e., the path which the packets are following in the Level 2 devices when going from the source to the destination. l2trace feature will not support multicast traffic. The Level 2 path is of critical interest to the network operators because it represents the reality of what physical boxes are involved in moving the packets from the source to the destination. This feature will save administrators from having to look at STP ports (which is blocking/forwarding), lengthy cam table entries, corresponding port information and telnetting from one switch to another and doing the same thing over and over again. L2trace feature will show the L2 path taken by the packet by displaying the device and port information of all the intermediate Catalyst switches the packet traverses. The l2trace information for the intermediate devices will consist of the following:

- Device name
- Device type
- Device IP address
- In port name
- Out port name
- In port speed
- Out port speed
- In port duplex
- Out port duplex

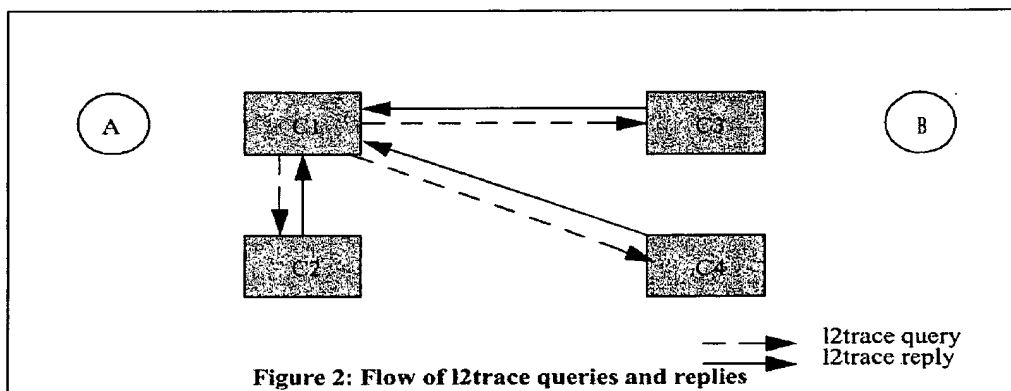
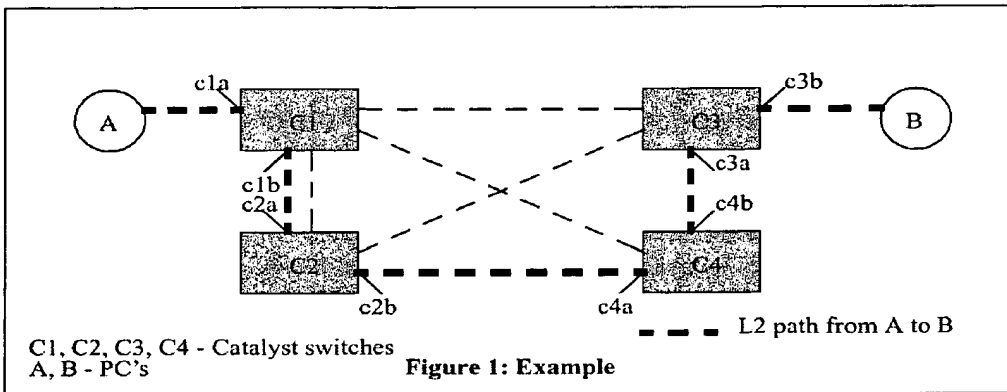
Section 5 shows the format of the information displayed.

3.0 Feature Description and Implementation

Obtaining the Level 2 path that a packet takes cannot be handled in similar lines as Level 3 traceroute, as the switching/forwarding of the packet is done in hardware using the forwarding engine (EARL). Instead, the following scheme will be used to get the Level 2 path a packet takes to reach from source to destination.

The following example will be used to illustrate the l2trace implementation.

Appl. No.: 09/824,844
Exhibit D (Page 3 of 6)



Setup: Catalyst switches C1, C2, C3 and C4 are physically connected as shown in the figure 1, by the dashed lines. Switches C1, C2, C3 and C4 are running SW of Orlando or later release. A and B are hosts (PC's) which are in the same subnet. Packets from A destined to B reach B without passing any L3 device in between. Host A is connected to C1 on port c1a and Host B is connected to C3 on port c3b. Let us say, in the above scenario, the packets from A destined to reach B, follow the path - A->c1a->c1b->c2a->c2b->c4a->c4b->c3a->c3b->B, as shown by the bold dashed line in Figure 1. That is, C1 forwards the packet from c1a->c1b, C2 forwards it from c2a to c2b, C4 forwards it from c4a to c4b and C3 forwards it from c3a to c3b. The goal of the l2trace feature is to provide the information of this complete path to the user which will include the port information (speed/duplex/FEC/GEC etc) of c1a,

Appl. No.: 09/824,844 Exhibit D (Page 4 of 6)
--

c2b, c2a, c2b, c3a, c3b, c4a and c4b, and the device information of C1, C2, C3, C4 (system name, device type and device IP address).

EARL table, CDP database and the IP stack on the switch will be used to derive the above l2trace information.

step 1: User requests for a 'l2trace' from mac address of A to mac address of B on the Catalyst switch C1.

step 2: C1 looks up its EARL table and identifies the destination port and vlan for A-mac. Once A-mac's destination port and vlan are identified (let us say, c1a and vl), C1 searches its EARL table for the destination port for B-mac. If C1 identifies the destination port (c1b) for B-mac in the same vlan as A-mac, i.e., vl, l2trace continues. Otherwise l2trace will be aborted with appropriate error message.

step 3: C1 searches the CDP database and identifies the IP address of the device (Catalyst switch) connected to c1b (IP address of C2). If C1 cannot detect the neighbor IP address on c1b or if the device connected to c1b is not a Catalyst switch of the series C5000 or C6000, l2trace will be aborted with appropriate error message.

step 4: C1 builds a l2trace query packet with all the relevant query information (like source MAC address, destination MAC address, vlan etc) and sends the l2trace query packet to C2 using IP. The transport mechanism used will be UDP. L2trace query will be destined to the IP address of C2 and will be sent to a predecided l2trace UDP port number.

step 5: C2 upon receiving the IP query, will first validate the query. It will then process the query in lines mentioned in step2 and step3 and identify IP address of the neighbor device connected to port c2b (IP address of C4), and send a reply back to C1 with device information of C2, port information of c2a, c2b and the IP address of C4.

step 6: C1 upon receiving a reply from C2 validates the reply and displays the L2 path information taken by the packet in the device C2 and sends the next l2trace query to C4 as in step 4.

step 7: C4 repeats step 5 and sends a response back to C1 with device and port information of C4 along with the IP address of the next device which the packet will traverse, C3.

step 8: C1 upon receiving a reply from C4 validates the reply and displays the L2 path information taken by the packet in the device C4 and sends the next l2trace query to C3 as in step 4.

step 9: C3 repeats step 5 and sends a response back to C1 with the device and port information of C3. Since the port c3b is directly connected to the device B, there will no CDP neighbor detected on port c3b. C3 will send a reply to C1 with the port and device information of the L2 path and also specify the next IP address as 0.

step 10: C1 repeats step 6. IP address of 0 in the l2trace reply from C3 to C1 indicates to C1 that there are no further devices to be traced. C1 after displaying the reply information from C3 terminates the l2trace process.

Figure 2, shows the flow of l2trace queries and replies between the devices. It should be noted that the L2 path is obtained by looking at the EARL tables in the Catalyst switches and not by actually sending a packet from A to B and tracing the path that the packet takes.

Appl. No.: 09/824,844 Exhibit D (Page 5 of 6)
--

5.0 User Interface

The **l2trace** feature will be made available to the user as the following new normal mode command.

- **l2trace <source> <destination> [vlan] [detail]**

User can use the above command to display the L2 path taken by the packets which start at the source and reach the destination. User can specify the source and destination addresses as either IP address or MAC address. When the source and destination belong to multiple vlans, user can specify the vlan number to be used in L2 path determination. User can specify the 'detail' option, when the user wishes to see the detailed l2trace output.

Examples:

Console> (enable) l2trace help

Usage: l2trace <src-mac-addr> <dest-mac-addr> [vlan] [detail]

l2trace <src-ip-addr> <dest-ip-addr> [detail]

(src-ip-addr, dest-ip-addr: IP alias or IP address)

Examples:

Appl. No.: 09/824,844
Exhibit D (Page 6 of 6)

Case 1: Source and destination in MAC address format. Vlan not specified. Detailed output

Console> (enable) l2trace 00-01-22-33-44-55 10-22-33-44-55-66 detail
l2trace vlan number is 10.

00-01-22-33-44-55 found in C5500 named wiring-1 on port 4/1 10Mb half duplex
C5500: wiring-1: 192.168.242.10: 4/1 10Mb half duplex -> 5/2 100MB full duplex
C5000: backup-wiring-1: 192.168.242.20: 1/1 100Mb full duplex -> 3/1-4 FEC attached
C5000: backup-core-1: 192.168.242.30: 4/1-4 FEC attached -> 1/1-2 GEC attached
C6000: core-1: 192.168.242.40: 1/1-2 GEC attached -> 2/1 10MB half duplex.
10-22-33-44-55-66 found in C6000 named core-1 on port 2/1 10MB half duplex.

Case 2: Source and destination in MAC address format. Vlan not specified. Default output

Console> (enable) l2trace 00-01-22-33-44-55 10-22-33-44-55-66
4/1 : 192.168.242.10 : 5/2
1/1 : 192.168.242.20 : 3/1-4
4/1-4 : 192.168.242.30 : 1/1-2
1/1-2 : 192.168.242.40 : 2/1

Case 3: Source and destination address in IP format.

Console> (enable) l2trace user-1-pc user-2-pc detail
Mapping IP address to MAC Address
user-1-pc -> 00-01-22-33-44-55
user-2-pc -> 10-22-33-44-55-66
l2trace vlan number is 10

00-01-22-33-44-55 found in C5500 named wiring-1 on port 4/1 10Mb half duplex
C5500: wiring-1: 192.168.242.10: 4/1 10Mb half duplex -> 5/2 100MB full duplex
C5000: backup-wiring-1: 192.168.242.20: 1/1 100Mb full duplex -> 3/1-4 FEC attached
C5000: backup-core-1: 192.168.242.30: 4/1-4 FEC attached -> 1/1-2 GEC attached
C6000: core-1: 192.168.242.40: 1/1-2 GEC attached -> 2/1 10MB half duplex.
10-22-33-44-55-66 found in C6000 named core-1 on port 2/1 10MB half duplex.

Case 4: Source is not directly attached to the switch

Console> (enable) l2trace 1.2.3.4 5.6.7.8 detail
Doing IP to MAC lookup
1.2.3.4 -> 00-01-22-33-44-55
5.6.7.8 -> 10-22-33-44-55-66
l2trace vlan number is 10.
Attention: Source 00-01-22-33-44-55 is not directly attached to this system. It is attached to C5500 name wiring-1.

00-01-22-33-44-55 found in C5500 named wiring-1 on port 4/1 10Mb half duplex
C5500: wiring-1: 192.168.242.10: 4/1 10Mb half duplex -> 5/2 100MB full duplex
C5000: backup-wiring-1: 192.168.242.20: 1/1 100Mb full duplex -> 3/1 FEC attached
C5000: backup-core-1: 192.168.242.30: 4/1 FEC attached -> 1/1 GEC attached
C6000: core-1: 192.168.242.40: 1/1 GEC attached -> 2/1 100 Mb full duplex.
10-22-33-44-55-66 found in C6000 named core-1 on port 2/1 100 Mb full duplex.

BEST AVAILABLE COPY